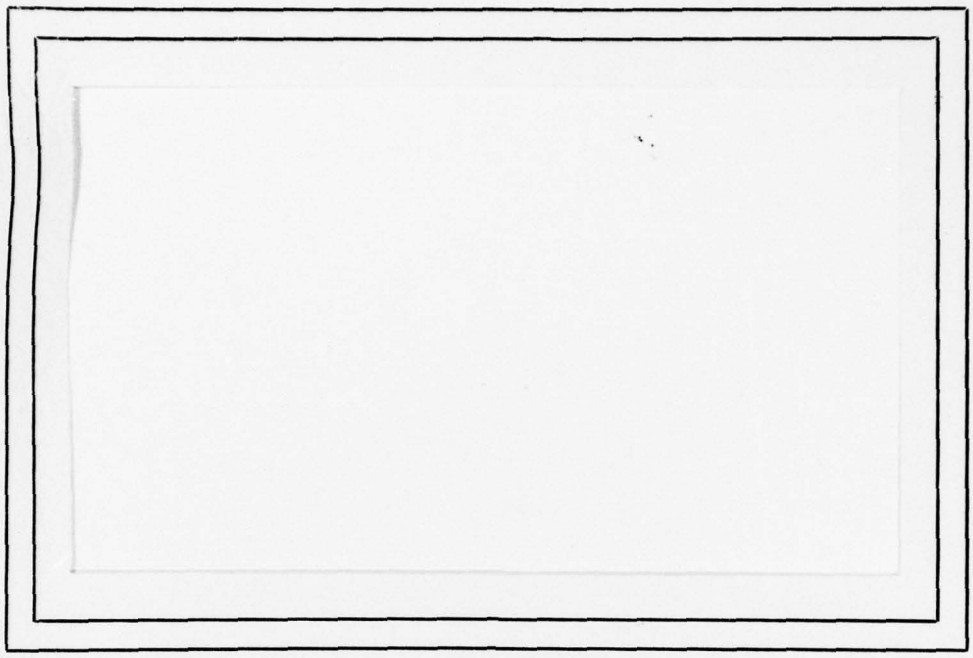
AD-A078 369 MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER F/G 5/8
DECISION RULES FOR CHOICE OF NEIGHBORS IN RANDOM FIELD MODELS O--ETC(U)
AUG 79 R CHELLAPPA , R L KASHYAP , N AHUJA AFOSR-77-3271
UNCLASSIFIED TR-802 AFOSR-TR-79-1271 NL

1 OF 1
AD A078369

END
DATE FILMED
1-80
DDC

LEVEL II



COMPUTER SCIENCE
TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

LEVEL II

(11)

TR-802
AFOSR-77-3271

August 1979

DECISION RULES FOR CHOICE OF NEIGHBORS IN RANDOM FIELD MODELS OF IMAGES

R. Chellappa[1,2]
R. L. Kashyap[2]
N. Ahuja[1*]

ABSTRACT

Consider the parametric representations of a homogeneous (not necessarily isotropic) random field by a set of neighborhood models. Using Bayesian methods and spectral representation of random fields, this paper develops decision rules for choosing a model from a set of r such random field models which differ in their neighborhoods. The theory developed here will find applications in image modeling and texture characterization.

The support of the U.S. Air Force Office of Scientific Research under Grant AFOSR-77-3271 is gratefully acknowledged, as is the help of Kathryn Riley in preparing this paper. The first author is grateful to Prof. Azriel Rosenfeld for his continued support, encouragement, and helpful discussions.

---

1. Computer Science Center, University of Maryland, College Park, MD.
2. School of Electrical Engineering, Purdue University, W. Lafayette, IN.
* Present address: Coordinated Science Laboratory, University of Illinois, Urbana, IL.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(9) Technical rept.

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFOSR-TR-79-1271	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): DECISION RULES FOR CHOICE OF NEIGHBORS IN RANDOM FIELD MODELS OF IMAGES	5. TYPE OF REPORT & PERIOD COVERED: Interim
	6. PERFORMING ORG. REPORT NUMBER: TR-802
7. AUTHOR(s): R. Chellappa, R. L. Kashyap, N. Ahuja	8. CONTRACT OR GRANT NUMBER(s): AFOSR-77-3271
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Computer Science Center, University of Maryland, College Park, MD 20742	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F; 2304/A2
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332	12. REPORT DATE: August 1979 13. NUMBER OF PAGES: 37
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image processing
Pattern recognition
Image models
Random fields

403 018 JOB

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Consider the parametric representations of a homogeneous (not necessarily isotropic) random field by a set of neighborhood models. Using Bayesian methods and spectral representation of random fields, this paper develops decision rules for choosing a model from a set of r such random field models which differ in their neighborhoods. The theory developed here will find applications in image modeling and texture characterization.

Unclassified

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE

1. Introduction

We are interested in finding an adequate model for a scalar observation set $\{ y(\underset{\sim}{s}), \underset{\sim}{s} = (s_1, s_2) \in \Omega_S \}$ obtained from a homogeneous Gaussian random field, where Ω_S is a square grid of side N_1, so that $1 \leq s_i \leq N_1$, $i = 1,2$. The random field is not necessarily isotropic. Such problems are of interest in various applications such as plant modeling [1], image restoration [2-3], seismology [4], and image modeling [5-6].

In situations such as these, any observation at (i,j) is statistically dependent on the observations over a neighborhood of (i,j), in contrast with the familiar univariate time series models, where the observation at any instant is dependent only on the past observations. Neighborhood models seem to be more appropriate for images since for an image there is no essential difference between the neighbors on one side and those on the other. Our prime interest in this paper is in developing decision rules for choosing an appropriate neighborhood from among a given set of neighborhoods. For instance, the set of north, south, east and west neighbors constitutes a particular choice of neighborhood.

The main approaches to choosing neighbors in model building are the following:

1) Pairwise testing

2) Akaike's information criterion (AIC)

3) The Bayesian approach

The reader is referred to the statistical literature [7] for a discussion of pairwise testing. The main criticisms of this approach are that the resulting decision rules are not consistent, i.e., the probability of choosing an incorrect model does not go to zero even as the number of observations goes to infinity. Also, the decision rules are not transitive, i.e., given three hypotheses C_1, C_2, and C_3, if C_1 is preferred to C_2 and C_2 is preferred to C_3, then it does not follow that C_1 is preferred to C_3 [8].

The AIC method [9] considers the so-called AIC statistics of the given observations for each model. The best model is the one which minimizes the AIC statistic. This method gives transitive but not consistent decision rules [10].

In the Bayesian approach [8] of fitting models to data, various possible models are postulated as mutually exclusive hypotheses C_i, $1 \leq i \leq r$. The hypothesis that maximizes the posterior probability density $P(C_i | y(\underset{\sim}{s}), \underset{\sim}{s} \in \Omega_S)$ is chosen as the correct model with minimum probability of error. This approach involves obtaining an expression for the likelihood of the observations and integrating it over the parameters using an appropriate prior probability density function.

In this paper we propose a Bayesian method for finding a neighborhood model for a random field. We take a transform domain approach using the spectral representation for the random field. Specifically, using the asymptotic Gaussian properties of the finite Fourier transforms $\{Z(\underset{\sim}{\lambda}), \underset{\sim}{\lambda} \in \Omega_\lambda\}$ where Ω_λ is a set of discrete frequencies $\underset{\sim}{\lambda} = (\lambda_1, \lambda_2)$ with components $\lambda_i = 2\pi k_i/N_1$, $1 \leq k_i \leq N_1$, $i = 1,2$, an explicit expression is given for $p(Z(\underset{\sim}{\lambda}), \underset{\sim}{\lambda} \in \Omega_\lambda | \underset{\sim}{\theta}_k, C_k)$, the dependence on the parameters appearing through the spectral density function of the field.

We integrate this probability density function w.r.t. arbitrary but regular prior probability density functions $p(\underset{\sim}{\theta}_k | C_k)$ to obtain $p(Z(\underset{\sim}{\lambda}), \underset{\sim}{\lambda} \in \Omega_\lambda | C_k)$. Using this expression and the prior probabilities $P(C_k)$, $1 \leq k \leq r$ of the hypothesis, a decision rule for choosing a model with minimum probability of error is designed.

The usual criticisms against the use of prior densities are answered by showing that the contributions of these terms are of order O(1) and hence asymptotically insignificant. A decision rule suppressing the terms involving the prior densities is also given. Though this rule does not have the minimum error rate property, it is asymptotically consistent.

The theory developed here has also been considered in an earlier report [5] motivating the use of statistical inference

theory for image modeling. The transform domain approach used in this paper can be easily extended to include moving average and autoregressive moving average models in a unified representation, viz., the spectral density function. Various special cases of interest, including random fields specified by one-sided models, and random processes represented by bilateral and unilateral models, are considered.

The organization of the paper is as follows: In Section 2, an explicit expression is derived for the probability density of the transforms of the observations given the neighborhood model obeyed by the observations. In Section 3, the problem of finding the appropriate neighborhood is posed as a class selection problem and decision rules are designed for choosing this neighborhood. Section 4 discusses the properties of the decision rules. Some special cases of the theory developed here are considered in Section 5. A brief discussion is given in Section 6, and some applications are considered in Section 7.

2. Preliminaries

A. Definitions and Notation : We are given a set of observations $\{y(\underset{\sim}{s}),\ \underset{\sim}{s} = (s_1,s_2) \in \Omega_S\}$, where Ω_S is a square grid of side N_1 and $1 \leq s_i \leq N_1$, $i = 1,2$. The random field is homogeneous and Gaussian but not necessarily isotropic. A stochastic field is said to be homogeneous if the following condition is satisfied:

$$R(s,t) = E(y(s) - E(y(s)))(y(t) - E(y(t)))$$
$$= R_1(s - t)$$

--i.e., the covariance function is translation invariant. In addition, if the covariance function R(s,t) is also invariant to rotation, the random field is called isotropic. In general, images are not isotropic and hence the random field models of interest to us are only homogeneous and not necessarily isotropic.

Consider the stochastic field $(y(\underset{\sim}{s}),\ \underset{\sim}{s} \in \Omega_S)$ satisfying

$$E(\underset{\sim}{\phi},\rho):\ y(\underset{\sim}{s}) + \sum_{k=1}^{m} \phi_k y(\underset{\sim}{s} + \underset{\sim}{q}_k) = \sqrt{\rho}\,\eta(\underset{\sim}{s}) \qquad (2.1)$$

$$\forall \underset{\sim}{s} \in \Omega_S,\ \underset{\sim}{q}_k \in Q,\ \forall k = 1,\ 2,\ldots m$$

$$Q = \{\underset{\sim}{q}_k = (q_{k1},q_{k2}),\ k = 1,2,\ldots m,\ \underset{\sim}{q}_k \neq (0,0),$$

$$q_{ki} \text{ are integers}\}$$

Here $\{u(\underset{\sim}{s}),\ \underset{\sim}{s} \in \Omega_S\}$ is a Gaussian I.I.D. sequence with zero mean and unit variance. In what follows, for simplicity, $\underset{\sim}{s}$ and $\underset{\sim}{q}_k$ will be denoted by s and q_k, dropping the vector notation.

Eq. (2.1) is characterized by an unknown (m+1) dimensional vector $\underset{\sim}{\theta}^T = (\underset{\sim}{\phi}^T, \rho)$ such that $\phi_j \neq 0$, $j = 1,2,\ldots m$ and $\rho > 0$. Eq. (2.1) represents the dependence of an observation at location (s_1, s_2) on its neighbors in the direction specified by Q. When q_{k1} and q_{k2} take only nonpositive values we obtain models where the observation at location (s_1, s_2) is a linear combination of the observations in a one-sided neighborhood. When s and q are scalars, (2.1) represents a one-dimensional autoregressive model. We assume that the coefficient vector $\underset{\sim}{\phi}$ in (2.1) is restricted to ensure homegeneity of the stochastic field.

The two-dimensional $\underset{\sim}{z}$ transform of (2.1) where $\underset{\sim}{z} = (z_1, z_2)$ is a complex vector given by

$$H(z_1, z_2, \underset{\sim}{\phi}) = \left[1 + \sum_{k=1}^{m} \phi_k z_1^{q_{k1}} z_2^{q_{k2}}\right]^{-1} \tag{2.2}$$

and the spectral density function evaluated at frequency $\underset{\sim}{\lambda} = (\lambda_1, \lambda_2)$ of the field is given by

$$S_y(\underset{\sim}{\lambda}, \underset{\sim}{\phi}, \rho) = \rho \left|\left| H(e^{j\lambda_1}, e^{j\lambda_2}, \underset{\sim}{\phi}) \right|\right|^2 \tag{2.3}$$

$$j = \sqrt{-1}, \quad \forall \underset{\sim}{\lambda} \in \Omega_\lambda$$

To make our notations clear, we consider a few examples. In what follows, we drop the vector notation ~ for λ.

Example 1: East, West, North, and South model:

$$Q = \{(1,0), (0,-1), (-1,0), (0,1)\}$$

The equation for $y(\cdot)$ is

$$y(s_1,s_2) + \phi_1 y(s_1+1,s_2) + \phi_2 y(s_1,s_2-1)$$
$$+ \phi_3 y(s_1-1,s_2) + \phi_4 y(s_1,s_2+1) = \sqrt{\rho}\, u(s_1 s_2)$$

the transfer function being given by

$$H(z_1,z_2,\underset{\sim}{\phi}) = [1 + \phi_1 z_1 + \phi_2 z_2^{-1} + \phi_3 z_1^{-1} + \phi_4 z_2]^{-1}$$

Example 2: One-sided models [11-13]

$Q = \{(0,-1), (-1,-1), (-1,0)\}$, the corresponding equation being

$$y(s_1,s_2) + \phi_1 y(s_1,s_2-1) + \phi_2 y(s_1-1,s_2-1)$$
$$+ \phi_3 y(s_1-1,s_2) = \sqrt{\rho}\, u(s_1,s_2)$$

The transfer function for this system is given by

$$H(z_1,z_2,\underset{\sim}{\phi}) = [1 + \phi_1 z_2^{-1} + \phi_2 z_1^{-1} z_2^{-1} + \phi_3 z_1^{-1}]^{-1}$$

Example 3: Time series models: When s and q are scalars and the set Q consists of positive and negative integers we obtain bilateral models [1]. When the set Q consists of negative integers, we obtain the familiar autoregressive models.

B. Expression for $p(Z(\lambda), \lambda \in \Omega_\lambda)$

In this section we are interested in deriving an explicit expression for the probability density of the transforms $(Z(\lambda), \lambda \in \Omega_\lambda)$ of the observations $(y(s), s \in \Omega_S)$, given that the observations obey the model in (2.1).

To this end, we first obtain an expression for $p(Z(\lambda), \lambda \in \Omega_\lambda | \underset{\sim}{\phi}, \rho)$ and then integrate over $(\underset{\sim}{\phi}, \rho)$ by using an arbitrary but otherwise regular prior probability density $p(\underset{\sim}{\phi}, \rho)$. An expression for $p(Z(\lambda), \lambda \in \Omega_\lambda | \underset{\sim}{\phi}, \rho)$ is obtained by using the stochastic properties of finite Fourier transforms [14-16]:

Theorem 1: Consider the finite Fourier transforms of the observations $(y(s), s \in \Omega_S)$ defined over a square grid $N_1 x N_1$. Denote the finite Fourier transform of $(y(s), s \in \Omega_S)$ by

$$Z(\lambda) = (N_1)^{-1} \sum_{s \in \Omega_S} e^{-j\lambda^T s} y(s), \quad \lambda \in \Omega_\lambda \qquad j = \sqrt{-1} \tag{2.4}$$

where Ω_λ is the set of discrete frequencies $\lambda = (\lambda_1, \lambda_2)$ with components $\lambda_i = 2\pi k_i / N_1$, $1 \leq k_1 \leq N_1$, $i = 1,2$. Then for an infinite observation field, the finite Fourier transforms are complex normally distributed with zero mean and independently at different frequences with variances

$$E(Z(\lambda) Z^*(\lambda)) = S_y(\lambda, \underset{\sim}{\phi}, \rho) \qquad \forall \lambda \in \Omega_\lambda \tag{2.5}$$

where

$$S_y(\lambda,\underset{\sim}{\phi},\rho) = \rho||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2 \tag{2.6}$$

in which

$||\alpha||$ denotes the modulus of the complex variable α.

This theorem involves two approximations; one involves the asymptotic uncorrelatedness of the random variables $(Z(\lambda),\ \lambda \in \Omega_\lambda)$ for different λ's, and the other involves equation (2.6). The smoother $S_y(\lambda,\underset{\sim}{\phi},\rho)$ is in the vicinity of the discrete frequency pairs $\lambda \in \Omega_\lambda$, the better is the approximation for moderate values of N_1. If the spectral density function $S_y(\lambda,\underset{\sim}{\phi},\rho)$ were constant, equality would hold in (2.6) for all values of N_1. Otherwise the distribution theory is exact only when the number of observations is infinite.

Using Theorem 1 an expression can immediately be written for the joint density of the finite Fourier transforms:

$$\begin{aligned}\ell n\ p(Z(\lambda),\ \lambda \in \Omega_\lambda|\underset{\sim}{\phi},\rho) &= -\frac{N}{2}\ell n 2\pi \\ &-\frac{1}{2}\sum_{\lambda\in\Omega_\lambda} \ell n\ S_y(\lambda,\underset{\sim}{\phi},\rho) \\ &-\frac{1}{2}\sum_{\lambda\in\Omega_\lambda} ||Z(\lambda)||^2/S_y(\lambda,\underset{\sim}{\phi},\rho)\end{aligned} \tag{2.7}$$

Substituting (2.6) in (2.7) we have

$$\begin{aligned}\ell n\ p(Z(\lambda),\ \lambda \in \Omega_\lambda|\underset{\sim}{\phi}\ \rho) &= -\frac{N}{2}\ell n 2\pi\rho \\ &-\frac{1}{2}\sum_{\lambda\in\Omega_\lambda} \ell n||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2 \\ &-\frac{1}{2\rho}\sum_{\lambda\in\Omega_\lambda} ||Z(\lambda)||^2/||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2\end{aligned} \tag{2.8}$$

It is interesting to compare the expression obtained here for the probability density of the transforms of observations and the expressions obtained in Whittle [1]. Whittle starts with an exact expression for the likelihood of the noisy variates $(u(s), s \in \Omega_S)$. Since for a general neighborhood model the Jacobian of the transformation from the noisy variates u(.) to the observations y(.) is difficult to evaluate, an approximate expression is used for the determinant. However, the expression obtained here is not an approximation to the likelihood function of the observations as in [1]. The desnity function considered here is the joint density of the finite Fourier transforms which is a one to one transformation with Jacobian unity (though a general proof can be given to establish this, a simple derivation is given in Appendix II for a 4x4 field). Consequently, the estimates of $\underset{\sim}{\phi},\rho$ obtained by maximizing $p(Z(\lambda), \lambda \in \Omega_\lambda | \underset{\sim}{\phi},\rho)$ are only approximately the maximum likelihood estimates. Before proceeding further we need the following assumption.

Assumption 1: The first and second derivations of $\sum_{\lambda \in \Omega_\lambda} \ell n ||H(e^{j\lambda_1}, e^{j\lambda_2}, \underset{\sim}{\phi})||^2$ and $\sum_{\lambda \in \Omega_\lambda} ||Z(\lambda)||^2 / ||H(e^{j\lambda_1}, e^{j\lambda_2}, \underset{\sim}{\phi})||^2$ w.r.t. $\underset{\sim}{\phi}$ exist for all $\underset{\sim}{\phi} \in R^m$.

To obtain $p(Z(\lambda), \lambda \in \Omega_\lambda)$ we integrate $p(Z(\lambda), \lambda \in \Omega_\lambda | \underset{\sim}{\phi},\rho)$ over $(\underset{\sim}{\phi},\rho)$ using an appropriate prior probability density

function. We do not make any specific assumption about the structure of the prior probability density functions. They need be regular but otherwise arbitrary. Letting $\underset{\sim}{\theta}^T = (\underset{\sim}{\phi}^T, \rho)$ and performing the integration

$$p(Z(\lambda),\ \lambda\in\Omega_\lambda) = \int\int p(Z(\lambda),\ \lambda\in\Omega_\lambda|\underset{\sim}{\phi},\rho)\, p(\underset{\sim}{\phi},\rho)\, d\underset{\sim}{\phi}\, d\rho \tag{2.9}$$

we arrive at

Theorem 2: As the rectangle of observations becomes large in all dimensions, the probability density $p(Z(\lambda),\lambda\in\Omega_\lambda)$ is given by

$$\ell n\ p(Z(\lambda),\ \lambda\in\Omega_\lambda) = -\frac{N}{2}\ell n\bar{\rho} - \frac{1}{2}\sum_{\lambda\in\Omega_\lambda}\ell n||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2 + \ell n\ p(\underset{\sim}{\bar{\theta}}) + (m+1)\ell n 2\pi - (m+1)\ell n\ N - a(N) - \frac{1}{2}\ell n[\det\nabla^2_{\theta_i\theta_j} G(\underset{\sim}{\theta})]_{\underset{\sim}{\theta}=\underset{\sim}{\bar{\theta}}} + O(1/N) \tag{2.10}$$

where

$$\underset{\sim}{\bar{\theta}}^T = (\underset{\sim}{\bar{\phi}}^T, \bar{\rho}) = \max_{\underset{\sim}{\phi}\in R^m,\ \rho>0} G(\underset{\sim}{\phi},\rho) \tag{2.11}$$

$$G(\underset{\sim}{\phi},\rho) = -[\ell n\rho + (1/N)\sum_{\lambda\in\Omega_\lambda}\ell n||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2 + (1/N)\sum_{\lambda\in\Omega_\lambda}||Z(\lambda)||^2/||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2] \tag{2.12}$$

and

$$a(N) = 0.5(N + N\ell n\ 2\pi) \tag{2.13}$$

and $O(1/N)$ denotes a deterministic constant term behaving like k_1/N for large N where k_1 is independent of N.

Theorem 2 gives an explicit expression for the probability density of the transforms of observations from a stochastic field characterized by the spectral density function $S_y(\lambda, \underset{\sim}{\phi}, \rho)$.

3. Decision rules for the choice of neighborhoods

For an image it is reasonable to assume that an observation at (i,j) will not significantly depend on distant neighbors. Hence we restrict our allowable neighborhood set to a maximum of eight neighbors, East, West, North, and South and the four diagonal neighbors. Thus our problem is to find an appropriate set of neighbors among the possible 2^8 neighborhood sets for the given image.

A. Definition of classes: We formulate the problem of choice of neighborhood as a class selection problem and derive optimal decision rules. A class is defined as a set of models having the same neighborhood set Q but differing in the parameters $\underset{\sim}{\phi}$ or ρ. The class C_i consists of models of the form

$$C_i: y(s) + \sum_{j=1}^{m_i} \phi_{ij} y(s + q_k) = \sqrt{\rho_i}\, u(s) \quad q_k \in Q_i,\ q_k \neq (0,0) \qquad (3.1)$$

such that $\phi_{ij} \neq 0$, $\rho_i > 0$, $j = 1,2,\ldots m_i$, $i = 1,2,\ldots r$, where r denotes the number of classes. Thus a class consists of an infinite number of models with the same neighborhood. The given set of observations $(y(s), s \in \Omega_S)$ is said to obey class C_i if $(y(s), s \in \Omega_S)$ obeys only one model in C_i. Two classes C_i and C_j are said to be mutually exclusive if the neighborhoods they represent differ in at least one neighbor.

Given r mutually exclusive classes (equivalently, distinct neighborhoods) C_i, i = 1,2,... r, and a set of observations (y(s), $s \in \Omega_S$), our aim is to find the most appropriate class for (y(s), $s \in \Omega_S$). The criterion for appropriateness is usually defined as a suitable criterion function and a decision rule for assigning (y(s), $s \in \Omega_S$) to one of the r classes is designed to minimize the criterion function. The criterion function can be chosen to reflect the particular needs of the problem, such as minimizing the average probability of error, etc.

B. Expression for $P(C_i|Z(\lambda),\ \lambda \in \Omega_\lambda)$

We first compute an expression for $P(C_i|Z(\lambda),\ \lambda \in \Omega_\lambda)$, the posterior probability density of the transforms of the given data having been generated by some model in C_i, for every i, i = 1,2,...r. Subsequently, we derive optimal decision rules to minimize the probability of error and discuss simplifications of the decision rules.

Let $\underset{\sim}{\theta}_i^T = (\underset{\sim}{\phi}_i^T, \rho_i)$, $\underset{\sim}{\phi}_i = (\phi_{i1}, \phi_{i2}, \ldots \phi_{im_i}$ and $p(\underset{\sim}{\theta}_i|C_i)$ be the prior probability density function of the parameters under class C_i. An expression for $P(C_i|Z(\lambda),\ \lambda \in \Omega_\lambda)$ is given by

Theorem 3: Let the observations obey the class C_i. Then the posterior probability density of the transforms of the observations is given by

$$\ell n P(C_i|Z(\lambda),\ \lambda \in \Omega_\lambda) = -\frac{N}{2}\ell n \bar{\rho}_i$$

$$- ((m_i+1)/2)\ell n\ N - \frac{1}{2}\sum_{\lambda \in \Omega_\lambda} \ell n \left|\left| H_i(e^{j\lambda_1}, e^{j\lambda_2}, \underset{\sim}{\bar{\phi}}_i) \right|\right|^2$$

$$+ \ell n\ p(\underset{\sim}{\bar{\phi}}_i, \bar{\rho}_i|C_i) - \frac{1}{2}\ell n[\det \nabla^2_{\theta_i\theta_j} G(\underset{\sim}{\theta})]_{\underset{\sim}{\theta}=\underset{\sim}{\bar{\theta}}} - a(N)$$

$$+ ((m_k+1)/2)\ell n\ 2\pi + \ell n\ P(C_i) - \ell n\ p(Z(\lambda),\ \lambda \in \Omega_\lambda) + O(1/N) \tag{3.2}$$

where

$$a(N) = 0.5(N + N\ell n 2\pi)$$

$$\underset{\sim}{\bar{\theta}}_i^T = (\underset{\sim}{\bar{\phi}}_i^T, \bar{\rho}_i) = \max_{\substack{\underset{\sim}{\phi}_i \in R^{m_i},\ \rho_i > 0}} \{G_i(\underset{\sim}{\phi}_i, \rho_i)\} \tag{3.3}$$

$$G_i(\underset{\sim}{\phi}_i,\rho_i) = -[\ell n\rho_i + (1/N) \sum_{\lambda\in\Omega_\lambda} \ell n||H_i(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi}_i)||^2$$

$$+ (1/N\rho) \sum_{\lambda\in\Omega_\lambda} ||Z(\lambda)||^2/||H_i(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi}_i)||^2], \tag{3.4}$$

$$p(Z(\lambda),\ \lambda\in\Omega_\lambda) = \sum_{k=1}^{r} p(Z(\lambda),\ \lambda\in\Omega_\lambda|C_k)P(C_k) \tag{3.5}$$

and

$P(C_i)$, $i = 1,2,\ldots r$ are the prior probabilities of the classes.

Proof: This follows from Theorem 2 and Bayes' rule,

$$P(C_i|Z(\lambda),\ \lambda\in\Omega_\lambda) = \frac{p(Z(\lambda),\ \lambda\in\Omega_\lambda|C_i)P(C_i)}{\sum_{k}^{r} p(Z(\lambda),\ \lambda\in\Omega_\lambda|C_k)P(C_k)} \tag{3.6}$$

C. Decision rules: Consider a 0-1 loss function L which assigns unit cost to a wrong assignment of classes and zero cost otherwise, i.e.

$$L[C_i, d(y(s), s\in\Omega_S) = C_j)] = 0 \qquad C_i = C_j$$
$$= 1 \qquad C_i \neq C_j \tag{3.7}$$

Since the finite Fourier transformation is one to one, the cost of wrong assignment of the observation set $(y(s), s\in \Omega_S)$ is the same as the cost of wrong assignment of the set $(Z(\lambda), \lambda\in\Omega_\lambda)$. Our intention is to choose the decision rule to minimize the risk function, i.e., minimize J(d), where

$$J(d) = \sum_{i=1}^{r} P(C_i) \int L[C_i, d(Z(\lambda), \lambda\in\Omega_\lambda)] \times$$
$$p(Z(\lambda), \lambda\in\Omega_\lambda | C_i) d|Z(\lambda), \lambda\in\Omega_\lambda| \tag{3.8}$$

Substituting the loss function in (3.7) we have

$$J(d(Z(\lambda), \lambda\in\Omega_\lambda) = C_j)$$
$$= \int \sum_{\substack{i=1 \\ i\neq j}}^{r} p\{C_i | Z(\lambda), \lambda\in\Omega_\lambda)\ p(Z(\lambda), \lambda\in\Omega_\lambda) d|Z(\lambda), \lambda\in\Omega_\lambda| \tag{3.9}$$

The optimal decision rule is

$$d^*(Z(\lambda), \lambda\in\Omega_\lambda) = C_k =$$
$$\text{argument } \sup_i \{P(C_i | Z(\lambda), \lambda\in\Omega_\lambda)\} \tag{3.10}$$

Substituting for the posterior probability function from Theorem 3 and omitting common terms we have the following decision rule:

Choose class C_i^* if

$$i^* = \arg\{\inf_i\{h_i(Z(\lambda),\ \lambda\in\Omega_\lambda\}\} \tag{3.11}$$

where

$$h_i(Z(\lambda),\ \lambda\in\Omega_\lambda) = \frac{N}{2}\,\ell n\bar{\rho}_i + m_i\,\ell n\,N$$

$$+\frac{1}{2}\sum_{\lambda\in\Omega_\lambda}\ell n\left|\left|H_i(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\bar{\phi}}_i)\right|\right|^2 + m_i\ell n\,2\pi$$

$$-\ell n\ p(\underset{\sim}{\bar{\theta}}_i|C_i) - \frac{1}{2}\ell n[\det\nabla^2_{\theta_i\theta_j}G(\underset{\sim}{\theta})]\Big|_{\underset{\sim}{\theta}=\underset{\sim}{\bar{\theta}}} - \ell n\ P(C_i) \tag{3.12}$$

D. Simplified Decision Rules: The decision rule given in (3.12) involves arbitrary quantities such as prior probability densities, which for the problem considered here are not known. Hence we suggest a decision rule in which the prior densities are suppressed. The decision rule no longer minimizes the average probability of error but can be shown to be asymptotically consistent using a proof similar to one found in [17].

The simplified decision rule is:

Choose class i* if

$$i^* = \arg\{\inf_i\{g_i(Z(\lambda),\ \lambda\in\Omega_\lambda\}\} \tag{3.13}$$

where

$$g_i(Z(\lambda),\ \lambda\in\Omega_\lambda) = N\ell n\bar{\rho}_i + \sum_{\lambda\in\Omega_\lambda} \ell n||H_i(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\bar{\Phi}}_i)||^2 + m_k\ell n\ N \tag{3.14}$$

Eq. (3.13) can be rewritten as

$$g_i(Z(\lambda),\ \lambda\in\Omega_\lambda) = N\ell n\ \rho_i^* + m_k\ell n\ N \tag{3.15}$$

where

$$\ell n\ \rho_i^* = \ell n\ \bar{\rho}_i + \frac{1}{N}\sum_{\lambda\in\Omega_\lambda} \ell n||H_i(e^{j\lambda_1},\ e^{j\lambda_2},\underset{\sim}{\bar{\Phi}}_i)\ \ ^2 \tag{3.16}$$

The form of the decision rule in (3.16) is characteristic of the Bayesian approach [8][17-19].

4. Properties of the decision rule

Asymptotic consistency: One of the important properties of a decision rule is the consistency property. A decision rule is said to be asymptotically consistent if the probability of choosing an incorrect model given the correct model goes to zero as the number of observations goes to infinity. We do not give an explicit proof for the consistency of the decision rule suggested in the previous section. A proof similar to that in [17] can be given to establish asymptotic consistency of the decision rule.

Generality: The theory developed here is valid for autoregressive spatial models. The theory can be easily extended to include moving average and autoregressive moving average models.

Parsimony: The expression for the decision function $g_k(z(\lambda), \lambda \in \Omega_\lambda)$ brings out the disadvantage of having too large a value for m_k. If we increase m_k, then ρ_i^* decreases, causing a decrease in $\ell n \rho_i^*$. Thus $N \ell n \rho_i^*$ and $m_k \ell n N$ balance each other. This illustrates the desirability of keeping the unknown parameters to a minimum.

Transitivity: The decisions regarding pairwise comparison of the classes are transitive. This is because the decision function $g_k(z(\lambda), \lambda \in \Omega_\lambda)$ does not depend on any parameter outside class C_i.

5. Special cases

In this section we consider two special cases of the general theory developed in the previous section.

Case (i): Random fields represented by one sided models:

For these cases [11-13], we have from an extension of a relation that is valid for a weakly stationary process

$$\ell n\rho = \frac{1}{4\pi^2} \int_{\pi}^{\pi}\int_{\pi}^{\pi} \ell n\, S_y(\lambda, \underset{\sim}{\phi}, \rho)\, d\lambda \tag{5.1}$$

Substituting for $S_y(\lambda, \underset{\sim}{\phi}, \rho)$ from (2.3), we have

$$\ell n\rho = \frac{1}{4\pi^2} \int_{\pi}^{\pi}\int_{\pi}^{\pi} (\ell n\rho - \ell n||H(e^{j\lambda}, \underset{\sim}{\phi})||^2)\, d\lambda \tag{5.2}$$

$$= \frac{1}{4\pi^2}\, 4\pi^2 \ell n\rho - \frac{1}{4\pi^2} \int_{\pi}^{\pi}\int_{\pi}^{\pi} \ell n||H(e^{j\lambda}, \underset{\sim}{\phi})||^2 d\lambda \tag{5.3}$$

Hence we have

$$\int_{\pi}^{\pi}\int_{\pi}^{\pi} \ell n||H(e^{j\lambda}, \underset{\sim}{\phi})||^2 d\lambda = 0 \tag{5.3}$$

Approximating the double integration by double summation, (5.3) reduces to

$$\sum_{\lambda \in \Omega_\lambda} \ell n||H(e^{j\lambda_1}, e^{j\lambda_2}, \underset{\sim}{\phi})||^2 = 0 \tag{5.4}$$

Using (5.4), the simplified decision rule becomes: Choose the model k* if

$$k^* = \text{argument}\{\min g_k(z(\lambda),\ \lambda \in \Omega_\lambda)$$

where

$$g_k(Z(\lambda),\ \lambda \in \Omega_\lambda)$$

$$= N\ell n\bar{\rho}_k + m_k \ell n\, N$$

Thus the main difference between the one-sided and neighborhood models is due to the term

$$\sum_{\lambda \in \Omega_\lambda} \ell n \left| \left| H_k(e^{j\lambda_1}, e^{j\lambda_2}, \underset{\sim}{\bar{\phi}}_k) \right| \right|^2$$

Case (ii): <u>One dimensional series represented by bilateral and unilateral methods</u> [1]:

For this case, s and q are all scalars and q takes positive and negative values. The theory developed for random fields carries through for this case analogously. By replacing the summation over fields by a single summation we obtain the results derived in [5].

6. Discussion

The inference of stationary Gaussian random fields has been previously considered by Whittle [1] and Larimore [4].

A useful entity to work with for statistical inference purposes is the likelihood of the observations. For models where the observation at (i,j) depends only on one sided neighborhoods, the Jacobian of the transformation from noise variates to observations is unity and hence the likelihood function can be easily written. However, for models where the observation at (i,j) depends on a neighborhood, the Jacobian of the aforementioned transformation is not unity and is difficult to evaluate. By approximating the Jacobian of the transformation Whittle obtains an approximate expression for the likelihood of observations of a random field. Likelihood ratio tests and significance levels have been used to identify a neighborhood. The decision rules using pairwise hypothesis tests are not consistent and transitive. Moreover, Whittle's procedure becomes complicated when models other than autoregressive are considered. Lastly, even for autoregressive models, the evaluation of the Jacobian is a nontrivial task.

Larimore [4] has reconsidered the problem of inference of random fields. The procedure developed in [4] uses spectral representation of the random field. Instead of the likelihood

of the observations, the probability density of the finite Fourier transforms is used for inference. As discussed in Section 2, the approximation involves uncorrelatedness and variance of the finite Fourier transforms. However, the term corresponding to the Jacobian of the transformation is easy to compute once the neighborhood is specified. Also, the AIC criterion has been used for model identification. It has been shown [10] that even in the one-dimensional case the AIC rule is inconsistent; i.e., the probability of choosing a higher order model than the true model is not zero even for the case of infinitely many observations. Consequently, the use of AIC rule for random field model identification is not desirable.

The theory developed here yields asymptotically (weakly) consistent decision rules for choosing neighborhoods of models. The simplified decision rule does not involve any arbitrary quantities such as prior densities.

One of the problems not considered in this paper is the computational aspects of the method. Maximization of the probability density function $p(Z(\lambda), \lambda \in \Omega_\lambda | \underset{\sim}{\phi}, \rho)$ could be difficult for the following reasons:

1) Starting values for the maximization algorithm are difficult to obtain.

2) The problem is really a constrained maximization, the constraints being laid by the homogeneity requirements.

3) Often the function has multiple maxima, and it is very important to find the absolute maxima. Additional work should be done in this area.

The expressions used for the probability density of the transforms of observations are approximate. Recently, an exact expression for the likelihood function of the observations from a homogeneous random field characterized by a parametric model as in (2.1) has been developed in [6]. This formulation lends itself to a theoretical treatment of many problems in the area of image processing, such as modeling, segmentation, etc.

7. Applications

The theory developed here will find applications in image modeling and texture characterization.

Earlier attempts at image modeling have used a one-dimensional analysis on a time series obtained by concatenation of successive rows [13][20]. This is clearly inadequate since intuition alone suggests that image models should be inherently two-dimensional. Recently, two-dimensional but one-sided neighborhood models for images have been suggested [21 - 22]. Since for images there exists no preferred direction, one-sided models are not sufficient to characterize images.

Recently, stochastic partial differential equations have been suggested for images in [2]. No effort has been made to statistically infer the particular model.

Efforts have been directed towards developing two-dimensional models in [23 - 24]. An observation at position (i,j) is assumed to statistically depend on neighbors on every side within a window of size (2M+1)(2N+1). Instead of using a truly two-dimensional procedure to infer the window size, the constants M and N, determined by using Akaike's FPE criterion for a one-dimensional, unilateral autoregressive model, are used to determine the window size.

In this paper we have proposed using the theory of

statistical inference to choose the neighborhood for an image. The choice of neighborhood models under consideration could be arbitrary as long as they are mutually exclusive. The models developed should be useful in coding [6],[13],[24], segmentation [6],[23] and restoration [3],[6] of images as well as in texture characterization.

APPENDIX I

We prove Theorem 2.

Consider equation (2.8), repeated below

$$\ell n\ p(Z(\lambda)\ \lambda\in\Omega_\lambda|\underset{\sim}{\phi},\rho)$$

$$= -\frac{N}{2}\ell n\ 2\pi\rho - \frac{1}{2}\sum_{\lambda\in\Omega_\lambda}[\ell n||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2$$

$$+ (1/\rho)||Z(\lambda)||^2/||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2] \qquad (1)$$

or

$$p(Z(\lambda),\ \lambda\in\Omega_\lambda|\underset{\sim}{\phi},\rho)$$

$$= (1/2\pi)^{N/2}\exp\{-\frac{N}{2}\ell n\rho - \frac{1}{2}\sum_{\lambda\in\Omega_\lambda}\ell n||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2 -$$

$$(1/2\rho)\sum_{\lambda\in\Omega_\lambda}||Z(\lambda)||^2/||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2\}$$

$$= (1/2\pi)^{N/2}\exp[\frac{N}{2}G(\underset{\sim}{\phi},\rho)] \qquad (2)$$

where

$$G(\underset{\sim}{\phi},\rho) = -[\ell n\rho + \frac{1}{N}\sum_{\lambda\in\Omega_\lambda}\ell n||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2$$

$$+ (1/N\rho)\sum_{\lambda\in\Omega_\lambda}||Z(\lambda)||^2/||H(e^{j\lambda_1},e^{j\lambda_2},\underset{\sim}{\phi})||^2]$$

Let $\underset{\sim}{\theta}^T = (\underset{\sim}{\phi}^T,\rho)$

To compute $p(Z(\lambda),\ \lambda\in\Omega_\lambda)$ we integrate $p(Z(\lambda),\ \lambda\in\Omega_\lambda|\underset{\sim}{\theta})$ over $\underset{\sim}{\theta}$ by using a prior probability density $p(\underset{\sim}{\theta})$:

$$p(Z(\lambda),\ \lambda\in\Omega_\lambda)$$

$$= \int p(Z(\lambda),\ \lambda\in\Omega_\lambda \mid \underset{\sim}{\theta})\ p(\underset{\sim}{\theta})\,d\underset{\sim}{\theta} \tag{4}$$

Substituting (2) in (4) we have

$$p(Z(\lambda),\ \lambda\in\Omega_\lambda)$$

$$= (1/2\pi)^{N/2}\int \exp[\tfrac{N}{2}\ G(\underset{\sim}{\theta})]p(\underset{\sim}{\theta})\,d\underset{\sim}{\theta} \tag{5}$$

Expanding $G(\underset{\sim}{\theta})$ in Taylor's series in $\underset{\sim}{\theta}$ about $\underset{\sim}{\theta} = \overline{\underset{\sim}{\theta}}$ where

$$\overline{\underset{\sim}{\theta}} = \max_{\underset{\sim}{\theta}}\ G(\underset{\sim}{\theta})$$

we have

LHS of (5)

$$= (1/2\pi)^{N/2}\int \exp(\tfrac{N}{2})[G(\overline{\underset{\sim}{\theta}}) + [\nabla_{\underset{\sim}{\theta}} G(\underset{\sim}{\theta})]^T_{\underset{\sim}{\theta}=\overline{\underset{\sim}{\theta}}}(\underset{\sim}{\theta}-\overline{\underset{\sim}{\theta}})$$

$$+ (\underset{\sim}{\theta}-\underset{\sim}{\theta})^T[\nabla^2_{\theta_i\theta_j} G(\underset{\sim}{\theta})]\Big|_{\underset{\sim}{\theta}=\overline{\underset{\sim}{\theta}}}(\underset{\sim}{\theta}-\overline{\underset{\sim}{\theta}}) + \ldots]p(\underset{\sim}{\theta})\,d\underset{\sim}{\theta} \tag{6}$$

By using the definition of $\overline{\underset{\sim}{\theta}}$ the linear term in the exponent vanishes. Thus the LHS of (6) reduces to

$$= (1/2\pi)^{N/2}\exp[(\tfrac{N}{2})G(\overline{\underset{\sim}{\theta}})]\ x$$

$$\int \exp[(\tfrac{N}{2})(\underset{\sim}{\theta}-\overline{\underset{\sim}{\theta}})^T[\nabla^2_{\theta_i\theta_j} G(\underset{\sim}{\theta})]\Big|_{\underset{\sim}{\theta}=\overline{\underset{\sim}{\theta}}}(\underset{\sim}{\theta}-\overline{\underset{\sim}{\theta}}) + \ldots]p(\underset{\sim}{\theta})\,d\underset{\sim}{\theta}$$

$$= (\tfrac{1}{2\pi})^{N/2}\exp[(\tfrac{N}{2})G(\overline{\underset{\sim}{\theta}})]p(\overline{\underset{\sim}{\theta}})\ (\tfrac{2\pi}{N})^{(m+1)/2}$$

$$x\ \frac{1}{\{\det[\nabla^2_{\theta_i\theta_j}(G(\underset{\sim}{\theta}))]_{\underset{\sim}{\theta}=\overline{\underset{\sim}{\theta}}}\}^{1/2}} + O(1/N) \tag{7}$$

Observing that

$$G(\underset{\sim}{\bar{\theta}}) = -[\ell n\bar{\rho} + (1/N) \sum_{\lambda \in \Omega_\lambda} \ell n||H(e^{j\lambda_1}, e^{j\lambda_2}, \underset{\sim}{\bar{\phi}})||^2] + 1] \tag{8}$$

we obtain Theorem 2.

APPENDIX II

Lemma: The Jacobian of the transformation from the observation set $(y(s),\ s\in\Omega_S)$ to the finite Fourier transform $(Z(\lambda),\ \lambda\in\Omega_\lambda)$ is unity.

Proof: For simplicity, we consider a 4x4 case. We have

$$Z(\lambda_1,\lambda_2) = (N_1)^{-1} \sum_{s\in\Omega_S} e^{-j(\lambda_1 s_1+\lambda_2 s_2)} y(s_1,s_2)$$

For a 4x4 case, $N_1 = 4$,

$$\Omega_S = \{i,j\} \qquad 1\le i,\ j\le 4$$

and $\Omega_\lambda = \{\frac{2\pi i}{4}\ ,\ \frac{2\pi j}{4}\}$, $1\le i,\ j\le 4$

In matrix notation,

$$\underset{\sim}{Z} = \underset{\sim}{J}\ \underset{\sim}{Y}$$

where $\underset{\sim}{Z}$ and $\underset{\sim}{Y}$ are vectors of finite Fourier transforms and observations arranged in lexicographic order. The matrix J (16x16) can be written as

$$\underset{\sim}{J} = (1/4)\begin{bmatrix} \underset{\sim}{A} & e^{-j\frac{\pi}{2}}\underset{\sim}{A} & e^{-j\frac{2\pi}{2}}\underset{\sim}{A} & e^{-j\frac{3\pi}{2}}\underset{\sim}{A} \\ e^{-j\frac{\pi}{2}}\underset{\sim}{A} & e^{-j\frac{3\pi}{2}}\underset{\sim}{A} & e^{-j\frac{\pi}{2}}\underset{\sim}{A} & e^{-j\frac{3\pi}{2}}\underset{\sim}{A} \\ e^{-j\frac{2\pi}{2}}\underset{\sim}{A} & e^{-j\frac{\pi}{2}}\underset{\sim}{A} & e^{-j2\pi}\underset{\sim}{A} & e^{-j\frac{3\pi}{2}}\underset{\sim}{A} \\ e^{-j\frac{3\pi}{2}}\underset{\sim}{A} & e^{-j\frac{3\pi}{2}}\underset{\sim}{A} & e^{-j\frac{3\pi}{2}}\underset{\sim}{A} & e^{-j\frac{3\pi}{2}}\underset{\sim}{A} \end{bmatrix}$$

where

$$\underset{\sim}{A} = \begin{bmatrix} e^{-j\pi} & e^{-j\frac{3\pi}{2}} & e^{-j2\pi} & e^{-j\frac{\pi}{2}} \\ e^{-j\frac{3\pi}{2}} & e^{-j\frac{\pi}{2}} & e^{-j\frac{3\pi}{2}} & e^{-j\frac{\pi}{2}} \\ e^{-j2\pi} & e^{-j\frac{3\pi}{2}} & e^{-j\pi} & e^{-j\frac{\pi}{2}} \\ e^{-j\frac{\pi}{2}} & e^{-j\frac{\pi}{2}} & e^{-j\frac{\pi}{2}} & e^{-j\frac{\pi}{2}} \end{bmatrix}$$

Using Kronecker product notation,

$$\underset{\sim}{J} = \underset{\sim}{B} \times \underset{\sim}{A}$$

where

$$\underset{\sim}{B} = e^{-j\frac{2\pi}{2}} \underset{\sim}{A}$$

Hence

$$\det \underset{\sim}{B} = \det \underset{\sim}{A}$$

From a Theorem regarding the characteristic roots of Kronecker products [25] the characteristic roots of $\underset{\sim}{A}x\underset{\sim}{B}$ are $a_i b_j$, where a_i are the characteristic roots of $\underset{\sim}{A}$ and b_j are the characteristic roots of $\underset{\sim}{B}$. Hence,

$$\det \underset{\sim}{J} = (1/4)^{16} \prod_{1 \le i,j \le 4} a_i b_j = (1/4)^{16} (\det A)^4 (\det B)^4$$

$$= 1$$

by direct evaluation of det A.

References

1. P. Whittle, "On stationary processes in the plane", Biometrika, Vol. 41, pp. 434-449, 1954.

2. A. K. Jain, "Partial differential equations and finite difference methods in image processing, part 1, image representation", Journal of Optimization Theory and Applications, Vol. 23, Sept. 1977, pp. 65-91.

3. A. K. Jain and J. R. Jain, "Partial difference equations and finite difference methods in image processing, part 2, image restoration", IEEE Trans. on Automatic Control, Vol. AC-23, pp. 817-833, Oct. 1978.

4. W. E. Larimore, "Statistical inference on stationary random fields", Proc. IEEE, Vol. 65, pp. 961-970, 1977.

5. R. Chellappa and N. Ahuja, "Statistical inference theory applied to image modeling", TR-745, Computer Science Center, University of Maryland, College Park, Md., 20742.

6. R. L. Kashyap,"Univariate and multivariate random field models for images", presented at Workshop on Image Modelling, Chicago, Ill., August 1979.

7. E. L. Lehmann, Testing Statistical Hypotheses, New York: Wiley, 1959.

8. R. L. Kashyap, "A Bayesian comparison of different classes of dynamic models using empirical data", IEEE Transactions on Automatic Control, Vol. AC-22, pp. 715-727, Oct. 1977.

9. H. Akaike, "A new look at statistical model identification", IEEE Trans. on Automatic Control, Vol. AC-19, pp. 716-723, 1974.

10. R. L. Kashyap, "Inconsistency of the AIC rule for time series models" (unpublished notes).

11. A. Habibi, "Two-dimensional Bayesian estimate of images", Proc. IEEE, Vol. 60, pp. 878-883, 1972.

12. D. P. Panda and A. C. Kak, "Recursive least squares smoothing of noise in images", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-25, pp. 520-524, Dec. 1977.

13. E. J. Delp, R. L. Kashyap, and O. R. Mitchell, "Image data compression using autoregressive time series models", Pattern Recognition (to appear).

14. D. R. Brillinger, Time Series Data Analysis and Theory, Holt, Rinehart and Winston, New York, 1974.

15. D. R. Brillinger, "Fourier analysis of stationary processes", Proc. IEEE, Vol. 62, pp. 1628-1642, 1974.

16. L. H. Koopmans, The Spectral Analysis of Time Series, Academic Press, New York, 1974.

17. R. L. Kashyap, "Optimal choice of AR and MA parts in autoregressive and moving average models" (unpublished notes).

18. H. Akaike, "A Bayesian analysis of the minimum AIC procedure", Ann. Inst. Statist. Math., Vol. 30, pp. 9-14, 1978.

19. G. Schwarz, "Estimating the dimensions of a model", Ann. Statistics, Vol. 6, pp. 461-464, 1978.

20. B. H. McCormick and S. N. Jayaramamurthy, "Time series model for texture synthesis", Intl. Journal of Computer and Information Sciences, Vol. 3, pp. 329-343, 1974.

21. J. T. Tou et al., "Pictorial texture analysis and synthesis", Proc. Third Intl. Joint Conf. on Pattern Recognition, 1976.

22. R. L. Kashyap, "Two-dimensional autoregressive models for images; parameter estimation and choice of neighbors", Proc. Workshop on Pattern Recognition and Artificial Intelligence, April 1978.

23. K. Deguchi and I. Morishita, "Texture characterization and texture-based image partitioning using two-dimensional linear estimation techniques", IEEE Trans. on Computers, Vol. C-27, August 1978.

24. K. Deguchi and I. Morishita, "Image coding and reconstruction by two-dimensional optimal linear estimation", Proc. 4th IJCPR, pp. 530-532, Nov. 1978.

25. R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 79-1271	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* DECISION RULES FOR CHOICE OF NEIGHBORS IN RANDOM FIELD MODELS OF IMAGES		5. TYPE OF REPORT & PERIOD COVERED Technical 6. PERFORMING ORG. REPORT NUMBER TR-802
7. AUTHOR(s) R. Chellappa R. L. Kashyap N. Ahuja		8. CONTRACT OR GRANT NUMBER(s) AFOSR-77-3271
9. PERFORMING ORGANIZATION NAME AND ADDRESS Computer Science Center University of Maryland College Park, MD 20742		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A2
11. CONTROLLING OFFICE NAME AND ADDRESS Math. & Info. Sciences, AFOSR/NM Bolling AFB Washington, DC 20332		12. REPORT DATE August 1979 13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Image processing
Pattern recognition
Image models
Random fields

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Consider the parametric representations of a homogeneous (not necessarily isotropic) random field by a set of neighborhood models. Using Bayesian methods and spectral representation of random fields, this paper develops decision rules for choosing a model from a set of r such random field models which differ in their neighborhoods. The theory developed here will find applications in image modeling and texture characterization.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*